

What is claimed is:

1. A computer implemented method for selecting probes for gene expression monitoring comprising:

- 5 a) predicting hybridization intensities of a plurality of candidate probes and their corresponding control probes; and
- b) selecting the candidate probe that has the highest intensity difference over its corresponding control probe.

10 2. The method of Claim 1 wherein said hybridization intensity (I) is determined using the equation:

$$I = \exp[\sum_i P_i S_i]$$

wherein said P_i is the value of the i th parameter and S_i is a value derived from the sequence of said probe.

15 3. The method of Claim 2 wherein said P_i is the free energy of a base in a given position of said sequence relative to a reference base selected from the group consisting of A, C, G and T.

20 4. The method of Claim 3 wherein said S_i is a functional of said sequence.

5. The method of Claim 3 wherein said P_i is determined empirically.

60
50
40
30
20
10
0

6. The method of Claim 5 wherein said P_i is determined by least square fit using a training data set.

5 7. The method of Claim 5 wherein P_i is determined using Cholesky decomposition.

8. The method of Claim 7 wherein said S_i for a given nucleotide type is replaced with a smooth function of probe base number.

10 9. The method of Claim 8 wherein said smooth function is a polynomial function.

10. The method of Claim 7 wherein said S_i for a given nucleotide type is replaced with a smooth function of the difference of a base to the central position.

15 11. The method of Claim 8 wherein said smooth function is a polynomial function.

12. The method of Claim 7 wherein said equation includes at least one parameter describing probe-probe hybridization.

20 13. The method of Claim 7 wherein said equation includes at least one parameter describing hairpin formation.

14. The method of Claim 7 wherein said equation includes parameters describing nearest neighbor interactions.

15. The method of Claim 7 wherein said P_i is determined using singular value decomposition.

16. A computer implemented method for predicting hybridization affinity of a probe against its target comprising calculating hybridization intensity (I) using the equation:

$$I = \exp\left[\sum_i P_i S_i\right]$$

17. wherein said P_i is the value of the i th parameter and S_i is a value derived from the sequence of said probe.

18. The method of Claim 16 wherein said P_i is the free energy of a base in a given position of said sequence relative to a reference base selected from the group consisting of A, C, G and T.

19. The method of Claim 17 wherein said S_i is a functional of said sequence.

20. The method of Claim 18 wherein said P_i is determined empirically.

21. The method of Claim 19 wherein said P_i is determined by least square fit using a training data set.

21. The method of Claim 20 wherein P_i is determined using Cholesky decomposition.

22. The method of Claim 21 wherein said S_i for a given nucleotide type is replaced
5 with a smooth function of probe base number.

23. The method of Claim 22 whereon said smooth function is a polynomial function.

24. The method of Claim 21 wherein said S_i for a given nucleotide type is replaced
10 with a smooth function of the difference of a base to the central position.
25. The method of Claim 24 wherein said smooth function is a polynomial function.
26. The method of Claim 21 wherein said equation includes at least one parameter
15 describing probe-probe hybridization.

27. The method of Claim 21 wherein said equation includes at least one parameter
describing hairpin formation.

20 28. The method of Claim 20 wherein said equation includes parameters describing
nearest neighbor interactions.

29. The method of Claim 28 wherein said P_i is determined using singular value decomposition.

30. The method of Claim 16 wherein said probe is immobilized on a solid substrate.

5

31. A computer software product for predicting hybridization affinity of a probe against its target comprising:

computer program code for calculating hybridization intensity (I) using the equation:

$$I = \exp\left[\sum_i P_i S_i\right]$$

wherein said P_i is the value of the i th parameter and S_i is a value derived from the sequence of said probe; and

a computer readable media for storing said computer program code.

32. The computer software product of Claim 31 wherein said P_i is the free energy of a base in a given position of said sequence relative to a reference base selected from the group consisting of A, C, G and T.

33. The computer software product of Claim 32 wherein said S_i is a functional of said sequence.

34. The computer software product of Claim 33 wherein said Pi is determined empirically.

35. The computer software product of Claim 34 wherein said Pi is determined by least
5 square fit using a training data set.

36. The computer software product of Claim 35 wherein Pi is determined using Cholesky decomposition.

37. The computer software product of Claim 36 wherein said Si for a given nucleotide
10 type is replaced with a smooth function of probe base number.

38. The computer software product of Claim 37 wherein said smooth function is a
polynomial function.

39. The computer software product of Claim 36 wherein said Si for a given nucleotide
15 type is replaced with a smooth function of the difference of a base to the central position.

40. The computer software product of Claim 39 wherein said smooth function is a
20 polynomial function.

41. The computer software product of Claim 36 wherein said equation includes at least one parameter describing probe-probe hybridization.

42. The computer software product of Claim 36 wherein said equation includes at least one parameter describing hairpin formation.

43. The computer software product of Claim 35 wherein said equation includes parameters describing nearest neighbor interactions.

44. The computer software product of Claim 18 wherein said P_i is determined using singular value decomposition.

45. A computer software product for selecting probes for gene expression monitoring comprising:

computer program code for predicting hybridization intensities of a plurality of candidate probes and their corresponding control probes;

computer program code for selecting the candidate probe that has the highest intensity difference over its corresponding control probe; and

a computer readable media for storing said codes.

46. The computer software product of Claim 45 wherein said computer code for predicting hybridization intensity (I) comprising code for calculating using the equation:

$$I = \exp\left[\sum_i P_i S_i\right]$$

wherein said P_i is the value of the i th parameter and S_i is a value derived from the sequence of said probe.

5 47. The computer software product of Claim 46 wherein said P_i is the free energy of a base in a given position of said sequence relative to a reference base selected from the group consisting of A, C, G and T.

48. The computer software product of Claim 47 wherein said S_i is a functional of said sequence.

49. The computer software product of Claim 47 wherein said P_i is determined empirically.

15 50. The computer software product of Claim 49 wherein said P_i is determined by least square fit using a training data set.

51. The computer software product of Claim 50 wherein P_i is determined using Cholesky decomposition.

20 52. The computer software product of Claim 51 wherein said S_i for a given nucleotide type is replaced with a smooth function of probe base number.

53. The computer software product of Claim 52 whereon said smooth function is a polynomial function.

5 54. The computer software product of Claim 52 wherein said S_i for a given nucleotide type is replaced with a smooth function of the difference of a base to the central position.

55. The computer software product of Claim 52 wherein said smooth function is a polynomial function.

56. The computer software product of Claim 51 wherein said equation includes at least one parameter describing probe-probe hybridization.

57. The computer software product of Claim 51 wherein said equation includes at least one parameter describing hairpin formation.

58. The computer software product of Claim 51 wherein said equation includes parameters describing nearest neighbor interactions.

20 59. The computer software product of Claim 58 wherein said P_i is determined using singular value decomposition.